

The Imperative for Data Curation

Joyce L. Ogburn

The processes of creation and expression of our scientific, social, and humanistic inspirations are culminating in a vast corpus of stunning and even life-changing documents, films, recordings, Web sites, and other media, including software. Advances in technology have enabled new kinds of scholarship—the most obvious and profound impact is occurring in the realm of science.

Science is an interwoven system of experimentation, observation, verification, and replication that demands access to durable research results. Investigations into scholarly communication and research practices have brought attention to the evolving conduct of science. Service opportunities have been revealed for supporting the research process, sustaining and capturing the non-published conversations of science, and curating the resulting data.¹ An ARL workshop in 2006 elicited many salient issues regarding data curation, and a new body of literature has been building.² Recent developments offer compelling reasons to engage in the future of data.

Since 2008, the National Institutes of Health (NIH) have mandated that researchers deposit their peer-reviewed, NIH-funded research articles in PubMed Central.³ The NIH already had in place a requirement for researchers to deposit their data with NIH and to do so in prescribed formats.⁴ Calls for stronger data management plans by other federal granting agencies are growing. Spurred by the National Science Foundation's (NSF) initiative to build a supportive infrastructure for science,⁵ campuses are forming committees and formulating Datatnet proposals that involve many segments of the institution, including libraries.⁶ In June 2009, Senators

Cornyn and Lieberman reintroduced the *Federal Research Public Access Act* that would direct other federal agencies to require the deposit of articles in a certified repository.⁷

These public investments in science are predicated on the idea that the sharing of research data and publishable results stimulates additional innovation and discoveries. Such an open system of knowledge demands an infrastructure that will endure well

Leaving digitally based information to languish in personal electronic filing drawers amid a jumble of unrelated information and with no plans for its survival guarantees its disappearance.

```
ERROR:
undefined
OFFENDING COMMAND:
excp
STACK:
1
0.0
-1
1
--nostringval--
--nostringval--
--nostringval--
--nostringval--
--nostringval--
-mark-
-mark-
-mark-
```